A Graph Theory-Based Automated Twin Recognition Technique for Electron Backscatter Diffraction Analysis

4 authors, including:

Cedric Pradalier
Georgia Institute of Technology
99 PUBLICATIONS   1,213 CITATIONS

SEE PROFILE

Laurent Capolungo
Georgia Institute of Technology
94 PUBLICATIONS   1,954 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Unstructured data mining and its applications to information systems and robotics. View project

Project   Long-term Natural Environment Monitoring for Lake and River Shores View project

A graph theory based automated twin recognition technique for Electron Back-scatter Diffraction analysis

C. PRADALIER[a], P.-A. JUAN[b], R. J. McCABE[c], L. CAPOLUNGO[c]

[a]*Georgia Tech Lorraine, UMI GT-CNRS 2958, France*
[b]*Sandia National Laboratories, Albuquerque, NM 87185, USA*
[c]*Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

Abstract

The present article introduces a new numerical framework and associated graphical user interface to perform high throughput microstructure statistical analysis from electron backscatter diffraction maps. Emphasis is placed on the detection of twin domains in hexagonal close-packed metals. Originally, the approach proposed leverages graph theory, group structures and associated numerical algorithms to automatically detect twins and unravel both their intrinsic characteristics features as well as those pertaining to their interactions. The proposed graphical interface allows for the detection and correction of unlikely twin/parent associations rendering the approach applicable to highly deformed microstructures. Twin statistics and microstructural data are classified and saved in a relational database that can be interrogated to reveal a wide spectrum of features of the microstructure. Illustration of the approach is performed in the case of Zirconium.

1. Introduction

Twinning is a prevalent deformation mode in hexagonal close-packed (hcp) materials [1–6]. From the geometrical standpoint the twinning transformation consists of a rotation of the crystal lattice of π radians [7–10] about the normal to the twinning plane or about the twinning shear direction. The former operation pertains to type I twins while the latter applies to type II twins. The twinning process is decomposed into three steps [11]. First, the twin nucleates in areas such as grain boundaries, crack tips, ledges or other interface defects where stresses are highly localized. As per atomistic simulations [12], the nucleation process occurs over a small activation volume -likely in the order of a few nm^3-. Then, the twin domain propagates transversely within the grain.Here too, atomistic simulations in correlation with transmission electron microscopy observations have revealed the intricate features of twin tips and of their transport processes. Among others, the role of interfacial steps (e.g. disclination dipoles, twinning disconnections) was suggested to be key. The last step consists of twin thickening: twinning partial dislocations/disconnections [13] glide along planes parallel to the twinning plane resulting in twin thickening. Stress levels required for nucleation and growth of twins may differ drastically [14].

Understanding the role of twinning on microstructure evolution and internal stress development entails that methodologies exist to predict -rather than fit- both nucleation and thickening processes. Such task

is particularly complex as the scales at which nucleation occurs is typically much smaller than that of the material point in which constitutive models are applied. Experimental observations of twinned microstructures are therefore critical. In particular, Electron back scattering diffraction (-hereafter EBSD) -which provides a voxelized mapping of the microstructure from the crystallographic orientation of each point- has historically been used to investigate the correlations between texture, loading conditions and the activation of twinning. While pioneering studies allowed for a relatively limited sampling of twinned microstructure -as the detection of twin domains was necessarily performed without use of computational methods-, current commercial EBSD softwares automatically extract key features of deformed microstructure (e.g. grain size etc.). However due to the complexity of twinned microstructures, in which several distinct and interacting twins belonging to different twinning modes can be simultaneously found in a given crystallite [6, 15], commercial solutions cannot yet automatically identify twin domains nor provide descriptive statistics of their geometry, interactions etc. In the past decade, a series authors have [15–18] first developed and then improved an automated EBSD analysis method capable of identifying twin and parent phases in materials with both hcp and orthorhombic crystal symmetry. The application of this method to the case of deformed Zr and Mg [19, 20] has allowed the generation of statistical revealing the correlations between grain orientation and variant selection, grain size and twin nucleation, etc. In turns this data has provided valuable guidance to modeling developments. Among others, probabilistic twin nucleation models have been proposed to predict twin variant selection consistently with experimental observations [19, 20]. This marks a departure from deterministic approaches [21–26].

The present study proposes a novel approach, Microstructure Evaluation Tool for Interface Statistics (METIS), to perform high throughput statistical analysis of twinned microstructures. The novel development introduced in what follows aims at allowing for the analysis of complex maps in which several twin modes are simultaneously active and in which twin-twin interactions are statistically relevant. In these complex scenarii, the use of a deterministic algorithm to detect twin domains and group them may lead to false identification. The idea proposed here is to develop a graphical user interface allowing for the assessment of the quality of the twin detection - and if necessary for its correction-. The study is structured as follows. Section 2 recalls the notations and conventions used to quantify disorientations between voxels and groups of voxels. Section 3 presents the details of the algorithm used to detect twins. In particular here, it is shown that the use of well established algorithms applying concepts anchored in group theory can successfully lead to the identification of twin domains. Section 4 demonstrates some of the most important features of the code. The case of deformed Zr is selected for this application.

2. Euler angles, quaternion rotation representations and their application to EBSD data

2.1. Euler angles and quaternion orientation and rotation representations

An EBSD map can be seen as an image, e.g. a square or hexagonal array of measurement points, where each measurement point (or pixel) gives the crystal local orientation as a set of Euler angles following the Z-X-Z convention (Bunge), denoted by (ϕ_1, Φ, ϕ_2). Crystal orientation can be obtained by applying a rotation

matrix, denoted R to the basic crystal structure. From a transformation perspective, the matrix R, also more explicitly denoted by R_c^w, corresponds to the transformation from the frame attached to the crystal to the world frame, also referred to as the sample frame. Conversely, the transformation from the world to the crystal will be denoted by R_w^c.

In materials science, use of often made of alternative means to represent rotations. Among others, both Rodrigues vector and unit quaternions are widely used. For the sake of understanding recall that the Rodrigues vector is a vector whose length is proportional to the amplitude of a given rotation and whose direction is the axis around which the rotation is applied. Similarly, quaternions allow for a compact representation of a rotation of angle θ around an axis v with 4 values (w, x, y, z) where $w = \cos\frac{\theta}{2}$ and $(x, y, z) = v \cdot \sin\frac{\theta}{2}$. By analogy with complex numbers, w is called the real part of the quaternion and v the imaginary part. When working with rotations, unit quaternions are preferred, i.e. $\sqrt{w^2 + x^2 + y^2 + z^2} = 1$. The advantage of quaternions lies in the existence of a multiplication operator allowing the preservation of the group structure while keeping the representation compact [27].

In formal terms coming from differential geometry, extracting the Rodrigues representation from a quaternion or a rotation matrix is referred to as using the logarithmic map of the differential manifold [28]. Recovering the quaternion from the Rodrigues representation is the exponential map. Therefore, the relationship between a given Rodrigues vector, u, and its equivalent quaternion, q, can be written as follows:

$$u = \log q \quad \text{and} \quad q = \exp u$$

With this, the amplitude of the rotation q, denoted by θ, can be expressed as the norm of $\log q$: $\theta = \|\log q\|$. The $SO(3)$ group is not a vector space. As a consequence, the usual norm of the Euclidean space $\mathbb{R}^3$, i.e. $\sqrt{x^2 + y^2 + z^2}$, does not apply. A more appropriate norm consists in the amplitude of the rotation, θ. Therefore, the norm of a rotation represented by its quaternion q is defined as

$$\|q\|_{so} = \|\log q\| \tag{1}$$

Although the norm of a rotation is not directly used for EBSD map analysis, the resulting distance leads to a unified definition of disorientation. Therefore, the distance between two rotations represented by q_1 and q_2 is denoted by:

$$d(q_1, q_2) = \left\|\log\left(q_1^{-1} \cdot q_2\right)\right\| = \left\|q_1^{-1} \cdot q_2\right\|_{so} \tag{2}$$

2.2. Application to EBSD data

The automated identification of twin domains relies on the computation of disorientation between between voxels, groups of voxels etc.. Consider the case of two measurements points, represented by their respective quaternions $q_1 = q_w^{c_1}$ and $q_2 = q_w^{c_2}$. The actual disorientation between these points consists of the rotation that needs to be applied to q_1 to transform it into q_2.

$$\delta(q_1, q_2) = q_2 \cdot q_1^{-1} = q_w^{c_2} \cdot (q_w^{c_1})^{-1} = q_w^{c_2} \cdot q_{c_1}^w = q_{c_1}^{c_2} \tag{3}$$

With this notation one notes that δ includes both the amplitude and the axis of the rotation. The latter property will be particularly useful to recognize and identify the mode and system of twins. Further, taking advantage of quaternion properties, the smallest positive rotation transforming q_1 into q_2 is identified and assigned to $\delta(q_1, q_2)$. For example, if the real part of $\delta(q_1, q_2)$ is negative, it implies that $|\frac{\theta}{2}| > \frac{\pi}{2}$ radians. In this case, $\delta(q_1, q_2)$ is then replaced by $-\delta(q_1, q_2)$ which corresponds to a rotation around the same axis but with an angle equal to $\theta + \pi$ radians. With this, the disorientation angle is smaller than π radians in absolute value. In addition, the disorientation can be assumed to be positive since a quaternion representing a rotation of an angle θ around a vector v is equal to the quaternion corresponding to a rotation of an angle $-\theta$ around $-v$.

Further, the disorientation measure, δ, must consider crystal symmetries. The hexagonal crystallographic structure is invariant by rotations around the c-axis by $k * \frac{\pi}{3}$, $k \in I\!N$ and by π radians around any vector lying in the basal plane. Quaternions associated with symmetries around the c-axis and vectors lying in the basal plane are denoted by $q_x(k)$ and $q_z(k)$, respectively, and are expressed as follows:

$$q_x(k) = \exp\left(k\pi\vec{x}\right) \tag{4}$$

$$q_z(k) = \exp\left(k\frac{\pi}{3}\vec{z}\right) \tag{5}$$

The set of possible disorientations between q_1 and q_2, denoted by $\Delta(q_1, q_2)$, is then defined as:

$$\Delta(q_1, q_2) = \{q_z(j) \cdot q_x(i) \cdot \delta(q_1, q_2)\}_{i=0\ldots1, j=0\ldots5} \tag{6}$$

In the case of h.c.p. materials, $\Delta(q_1, q_2)$ contains 12 elements. The definition of the disorientation quaternion, $Diso(q_1, q_2)$, and its norm, $\|Diso(q_1, q_2)\|_{so}$, results from the definition of $\Delta(q_1, q_2)$. Their expressions are, respectively:

$$Diso(q_1, q_2) = \underset{q \in \Delta(q_1, q_2)}{\arg\min} \|q\|_{so} \tag{7}$$

$$\|Diso(q_1, q_2)\|_{so} = \min_{q \in \Delta(q_1, q_2)} \|q\|_{so} \tag{8}$$

The symmetry around any vector lying in the basal plane allows the disorientation with an angle θ greater than $\pi/2$ radian to be equivalent to a disorientation with an angle equal to $\theta - \pi$, smaller in magnitude. If $\theta - \pi$ is negative, the negative sign is removed by considering the rotation of $-(\theta - \pi)$ around the opposite rotation vector. Symmetries also imply that the norm of the disorientation quaternion, $\|Diso(q_1, q_2)\|_{so}$, is always in the range of 0 to $\pi/2$ radian.
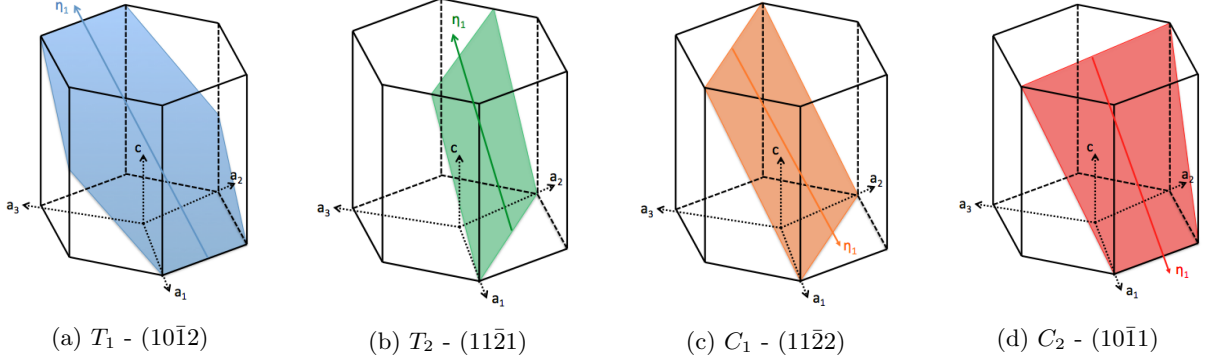
4

(a) T_1 - $(10\bar{1}2)$ (b) T_2 - $(11\bar{2}1)$ (c) C_1 - $(11\bar{2}2)$ (d) C_2 - $(10\bar{1}1)$

Figure 1: Schematic representation of twinning modes observed in Mg and Zr (T_2 and C_1 only appear in Zr). Twins are represented via their twinning planes, K_1.

2.3. Classification of twinning relationships

As per seminal work [1, 7, 10], in h.c.p metals a twin system is rigorously defined by the indices of its twinning plane, K_1 and of vector η_2 lying on a second invariant plane K_2. Alternatively and depending on the twinning mode, the second invariant plane, and the twinning shear direction, η_1, are used. Planes K_1 and K_2 and vectors η_1 and η_2 are all invariant to the twin transformation. When both K_1 and η_2 are rational, the twin mode is said to be of the first kind while a twin with rational indices for K_2 and η_1 is said to be of the second kind. The lattice reorientation induced by first and second kind twins respectively correspond to a rotation by 180° about either the normal to K_1 or to the twinning shear direction. In the case of compound twins, both of these rotations are equivalent. For the sake of generality the following four twin modes are considered; $\{10\bar{1}2\}$ tensile twinning or T_1, $\{11\bar{2}1\}$ tensile twinning or T_2, $\{11\bar{2}2\}$ compressive twinning or C_1 and $\{10\bar{1}1\}$ compressive twinning or C_2. Figure 1 shows a graphical representation of one of the six possible twin systems that can be activated for the four above mentioned twinning modes. All these twin modes are compound twins; their activation is material dependent. For example, T_1, T_2 and C_1 have been shown to be active in high-purity clock-rolled Zr samples subjected to loading along the through thickness and in-plane directions at 76K, [4, 29]. In Magnesium, only T_1 and C_2 are susceptible to nucleate and grow. Table 1 details the crystallography all twin modes cited in the above.

Twinning mode	Twinning plane, K_1	Twinning direction, η_1
T_1	$\{10\bar{1}2\}$	$<\bar{1}011>$
T_2	$\{11\bar{2}1\}$	$<\bar{1}\bar{1}26>$
C_1	$\{11\bar{2}2\}$	$<11\bar{2}\bar{3}>$
C_2	$\{10\bar{1}1\}$	$<10\bar{1}\bar{2}>$

Table 1: Twinning modes in Zr

With these definitions and accounting for the hexagonal structure considered here, disorientation quaternions can be indexed by k, such that $k \in [1, 6]$. These are written as follows:

$$q(k) = \exp\left[\pi \cdot \vec{\eta}(k)\right] \tag{9}$$

With,

$$\vec{\eta}(k) = \begin{bmatrix} \cos\left[\alpha(k)\right]\cos\left[\beta\right] \\ \sin\left[\alpha(k)\right]\cos\left[\beta\right] \\ \sin\left[\beta\right] \end{bmatrix} \tag{10}$$

where α denotes the angle between the first vector of the primitive hexagonal unit cell $\vec{a}_1$ and the vector $\vec{p}_s$ defined as the orthogonal projection on the basal plane of an upward pointing vector parallel to the twinning shear direction η_1. The angle β corresponds to the angle formed by the vector $\vec{p}_s$ and the twinning shear direction η_1. Note that the vector $\vec{p}_s$ and the orthogonal projection on the basal plane of the twinning shear direction η_1 are identical when dealing with tensile twinning modes while they are opposite in the case of compressive twinning. The values of both α and β associated with the six variants of each twinning mode are listed below:

- $\alpha(k) = \frac{(k-1)\pi}{3} - \frac{5\pi}{6}$ and $\beta = \arctan\left(\frac{\gamma}{\sqrt{3}}\right)$ for Tensile 1,

- $\alpha(k) = \frac{(k-3)\pi}{3}$ and $\beta = \arctan 2\gamma$ for Tensile 2,

- $\alpha(k) = \frac{k\pi}{3}$ and $\beta = \pi + \arctan\gamma$ for Compressive 1,

- $\alpha(k) = \frac{(k-1)\pi}{3} + \frac{\pi}{6}$ and $\beta = \pi + \arctan\left(\frac{2\gamma}{\sqrt{3}}\right)$ for Compressive 2.

With these definitions, identifying a given twin mode and system consists in finding the closest object to the disorientation $q_{c_1}^{c_2}$ existing between the parent and twin phases in $\mathcal{T}$, defined as the set of possible twinning relationships, within a given threshold $d_{\max}$:

$$\tau = \underset{t \in \mathcal{T}}{\arg\min}\, \delta\left(q_{c_1}^{c_2}, t\right) \quad \text{such that} \quad \left\| \delta\left(q_{c_1}^{c_2}, t\right) \right\|_{so} < d_{\max} \tag{11}$$

The set $\mathcal{T}$ contains 24 and 12 elements in the case of Zr and Mg, respectively.

3. Identification of grains, parent and twin phases

In EBSD scans, non-twinned grains, parent and twin phases are contiguous areas with a consistent orientation. Therefore, detecting grains and twins consists of identifying and grouping contiguous and consistent areas. These operations will be performed following a technique similar to the "super-pixels" technique [30, 31] based on graph theory and commonly used in image analysis. The whole process of twin recognition, parent phase identification, etc, relies on one tool of graph theory, that is used five times at different levels, which is the extraction of connected parts. Depending on the graph considered, connected parts can correspond to grains, twins or parent phases, etc. The extraction of connected parts is a well-known problem in graph theory for which efficient algorithms already exist [32] and for which the numerical

complexity increases linearly with the number of vertices belonging to the graph. In other words, the software identifies grains, twins and parent phases by generating five different graphs built on top of each other. Thus, the first graph aims at grouping all measurement points with consistent/similar orientation. The newly-formed groups can either be non-twinned grains or twin and parent phases in the case of twinned grains. The second graph is used to identify grains. The third and fourth graph enable the recognition of parent phases and the detection of higher-order twins, respectively. Finally, a fifth graph is built in order to group parts of a same twin into what will be referred to as a twin strip. The notions of graph, vertex, edge and connected part mentioned in the above are introduced in the following paragraph.

For the sake of demonstration, in what follows all maps, or portions thereof, correspond to experimental data from previous work by Kaschner et al. [4] on high-purity clock-rolled Zr samples loaded in compression along the through-thickness and one of the in-plane directions.

3.1. Segmentation of the EBSD map into fragments of consistent orientations
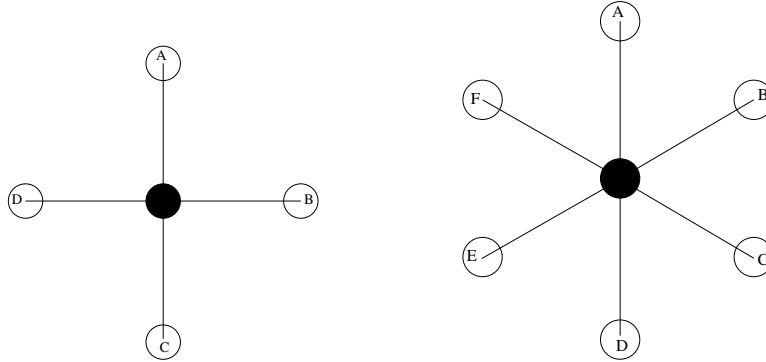

Figure 2: Example of neighboring relationships encountered in EBSD data. On the left, when measurement points are taken on a square grid, the pixel represented by the black disc has 4 neighbors represented by the white circles. On the right, when measurement points are taken on an hexagonal grid, each measurement has 6 neighbors.

The first step in the EBSD analysis consists of grouping all measurement points of similar orientation in order to form a first graph of connected fragments. Remember that, in graph theory, a graph $\mathcal{G}$ is a pair of set V, E of vertices and edges. Two vertices are said connected when an edge links them. Vertices are usually designated by integers. It implies that the edge (i, j) connects vertex i to vertex j. A path between two vertices k and l corresponds to a sequence of edges and vertices allowing to reach k from l, and reciprocally. In our case, edges and paths are undirected. In addition, a subset $W \subset V$ is a connected part of $\mathcal{G}$ if, for all pairs of vertices (i, j), $i \in W$, $j \in W$, there exists a path in $\mathcal{G}$ between i and j. In the present case, every measurement point is considered as a vertex and an edge between two neighboring pixels is created if the disorientation between them is smaller than a given threshold, e.g. 5 degrees. The type of measurement grid, i.e. square or hexagonal (Fig. 2), does affect the construction of the graph but does not affect the extraction of connected parts. In the graphical user interface, we display the disorientation via the thickness of edges, by making them directly proportional to the following weight:

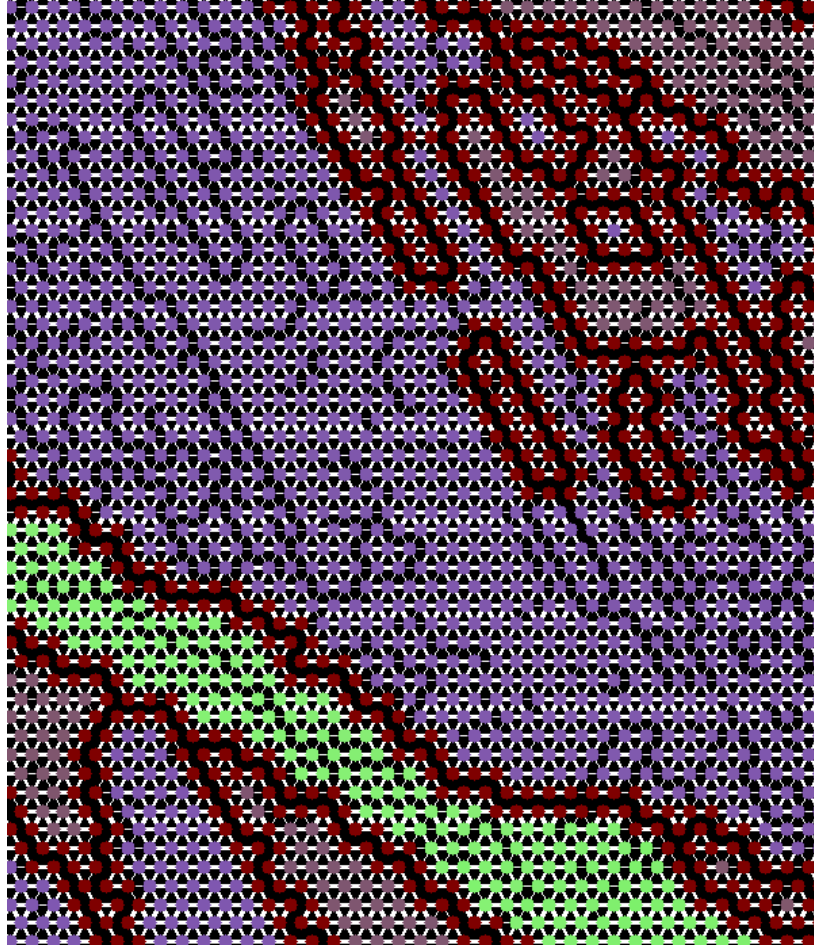$$w(q_1, q_2) = e^{\left(\frac{diso(q_1, q_2)}{L}\right)^2}$$

7

Figure 3: Graph grouping measurement points of consistent orientation in fragments. The colored circles correspond to EBSD measurement points, with the Euler angles mapped on the RGB cube, and the white lines are edges, whose thickness is proportional to the weight, w. Measurement points located along fragment borders are displayed in brown. Consequently, twins appear as areas delineated by a black border where the edge weight, w, becomes negligible.

where L is a threshold value. As a result, the smaller the disorientation between two pixels, the more strongly is displayed their mutual edge. The edge color indicates the nature of the disorientation. For example, an edge between two points of similar orientation appears in white. But, as shown in figure 4, disorientations corresponding to tensile 1 and compressive 1 twinning relationships are displayed in green and red, respectively.

In some cases, small groups of measurement points (e.g. three pixels or less) exhibit a too low resolution for the software to identify the fragment they belong to. These measurement points are then flagged as invalid pixels. However, a flood-filling algorithm was implemented to associate these points with the closest connected fragment [33].

3.2. Grouping of connected fragments into grains

The second step in the analysis of EBSD data is to group connected fragments of consistent orientation into grains. In a material free of twins or precipitates, this step is trivial since every connected fragment
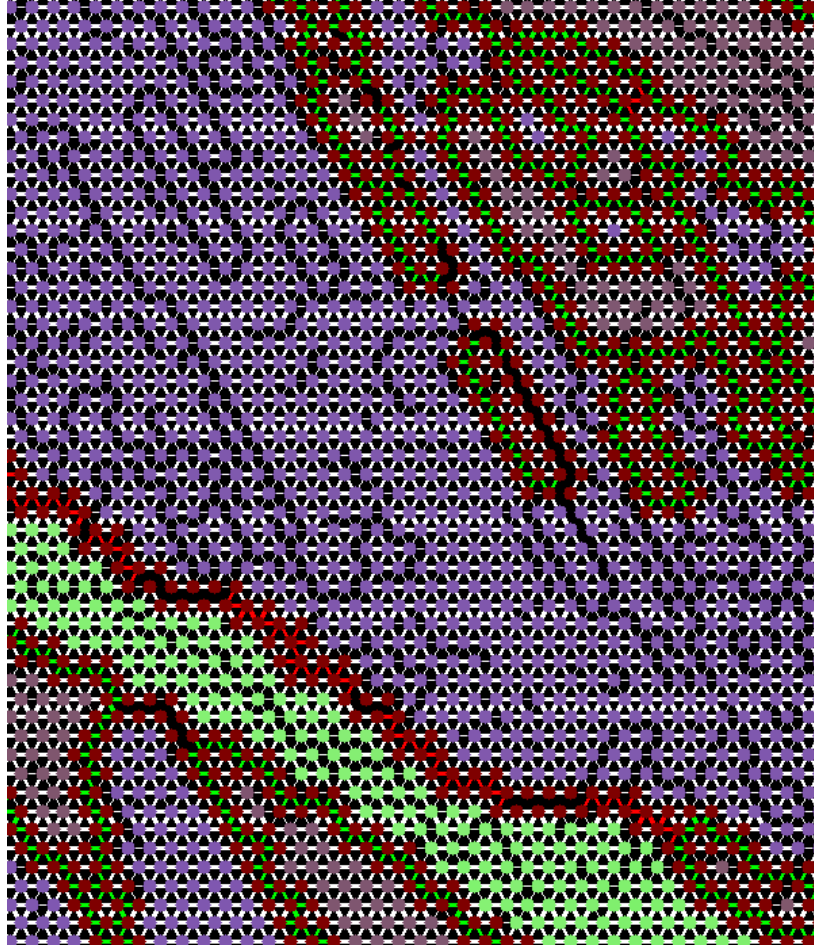
Figure 4: Graph grouping measurement points of consistent orientation in fragments with added twinning mode. Green and red edges linking border points, displayed in brown, indicate tensile and compressive twinning relations, respectively.

corresponds to a grain. However, when twinning occurs, different configurations have to be considered. Figure 5 depicts the 3 most typical twinning configurations observed in Zr scans. Therefore, a second graph, referred to as the twinning graph, is generated at the level of connected fragments to group them into grains. Vertices are now connected fragments and edges link two neighboring fragments for which the disorientation can be classified as a twinning relation.

The construction of such a graph relies on the measure of disorientation between two fragments of consistent orientation, i.e. connected parts, obtained after computing the average orientation of each fragment. Other hypotheses have also been considered such as, for example, the disorientation between the two centroids or the disorientation along the fragment boundary. Although these two methods would have been computationally less expensive, they present issues that overcome their benefits. First, the centroid might coincide with a bad measurement point or might be located outside of the fragment if the latter is non-convex. Second, the orientation angles obtained by EBSD for measurement points located near or along twin and grain boundaries are the least reliable. This explains why the choice has been made to use the average fragment orientation.

In addition, because $SO(3)$ is not an Euclidian space, the closed-form expression of the average is incorrect. Consequently, a specific algorithm, similar to the one computing average of quaternions, is used to determine the average orientation of connected EBSD measurements. Consider a set of n EBSD measurement points represented as quaternions q_i, $i = 1..n$. Assume that the initial average, m_0, is equal to q_1, i.e. $m_0 = q_1$. The average orientation of a fragment is estimated iteratively by computing the following two equations:

$$e_k = \frac{1}{n} \sum_{i=1}^{n} \log Diso(m_k, q_i)$$

$$m_{k+1} = m_k \cdot \exp e_k$$

The iteration stops when $\|e_k\|$ reaches a given threshold (5.10^{-4} in the present case). At this step, m_k corresponds to the best estimate of the fragment orientation.

Using the fragment orientation, it is now possible to build the twinning graph. The construction of this second graph and the extraction of its connected components enable the extraction of a significant amount of properties such as twin modes, twin systems, twin boundary lengths, the list of neighbors for each grain, the list of pixels belonging to grain joints, etc.
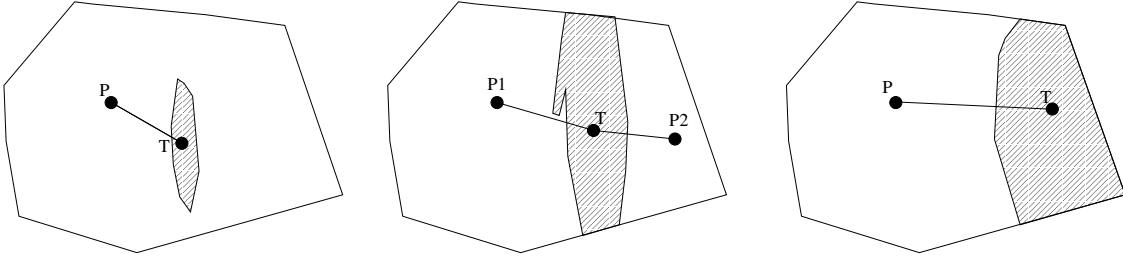

Figure 5: Three sample cases of twinning: on the left a single twin T in the middle of its parent P; in the middle a twin going across its parent and separating it into two fragments P1 and P2; on the right a grain appearing as two fragments next to each others.

3.3. Identification of parent phases

In twinned grains, parent phases are composed of one or several connected fragments, as shown in figure 5. Parent phases are then considered as sets of connected fragments of consistent orientation. To build such sets, the software generates a third graph over the EBSD map. Vertices are fragments and edges only link two fragments if they belong to the same grain and if they have a low respective disorientation. By construction, a connected component of this graph is a set of fragments embedded in the same grain with consistent orientation. By default, the parent phase is identified as the connected component of fragments occupying the largest part of the grain. In other words, the parent is the largest area with a consistent orientation.

In the graphical user interface, the user can intervene at this stage by disabling unlikely links between fragment or even by forcing the identification of the parent when a twin occupies a larger area than the

actual parent phase. Figures 6 and 7 show two parts of EBSD maps before and after edition of incorrect twinning relationships, respectively. Note that, as shown in figure 8, the software is capable of recovering complex grain structures.
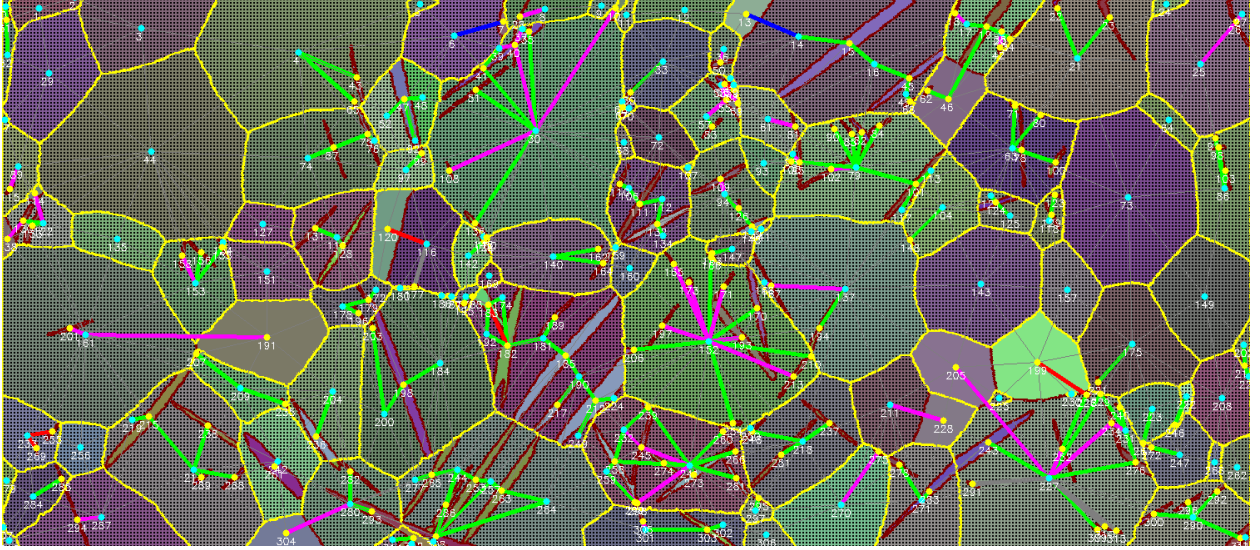

Figure 6: Automatic output for a Zr EBSD map. The sample is high-purity clock-rolled Zr plate loaded in compression along one of the in-plane directions up to 5% strain [29]. Yellow borders mark the grain joints, brown borders the twin joints. Green edges represent tensile 1 relation, magenta tensile 2, red compressive 1 and blue compressive 2.

3.4. Detection of higher order twins

A higher order twin is defined as a twin embedded in another. For example, a secondary or tertiary twin correspond to twins that nucleated in a primary and secondary twin, respectively. Depending on the material, the loading path and loading history, secondary twinning may occur. For example, it has been observed by Martin et al. [34] in Mg and appears on scans of Zr samples loaded along the through-thickness direction [4]. Tertiary twinning is more unlikely and statistically irrelevant. However, the software is still capable of identifying tertiary and higher order twins if necessary. The identification of these twins relies on the graph of connected fragments used to build the grains. In this graph, a twin of order n has an edge (i.e. an identifiable twinning relation) with a $(n-1)^{th}$-order twin or parent $(n-1=0)$, but no identifiable relation with $(n-k)^{th}$-order twins, with $k > 1$. An example of such a situation is seen in figure 9 where the orange secondary twin (marked with a blue dot in its center) is embedded in the gray twin (marked by a yellow dot). Although the orange twin shares a border with the parent phase, its disorientation with respect to the parent domain does not correspond to one of the previously detailed twinning relations.

The present definition of twinning order leads to a simple recursive implementation of higher-order twins. The initialization step consists of considering that all parent fragments previously identified are of twinning order 0. To find n-th order twins, all nodes, i.e. fragments, with a twinning order strictly lower than n are removed as well as all edges with one extremity corresponding to one of these nodes. The connected components of the resulting graph are twins of order n. If such a connected component contains several
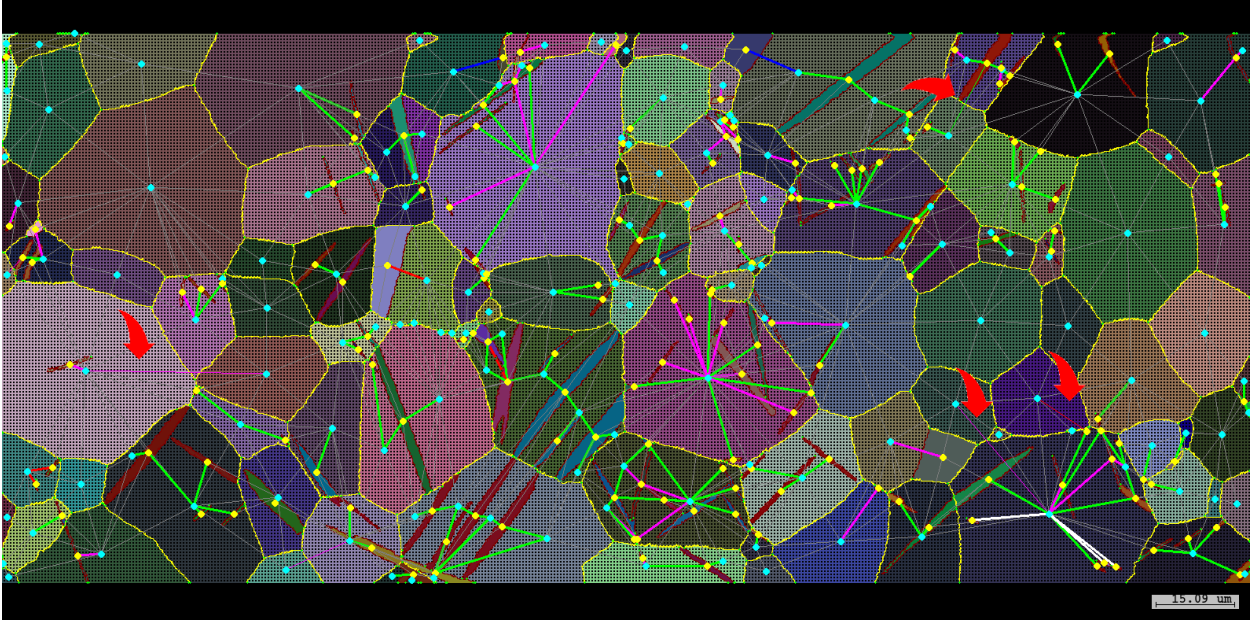
Figure 7: Same map as figure 6, but with manual edition of 4 incorrect links. The disabled links are displayed as thin edges and highlighted with the red arrow.

fragments, these fragments can be separated in groups of consistent orientation and identify the "parent" phase, i.e. n^{th}-order twin itself, as the largest one. All fragments not identified as the "parent" fragment correspond to twins of order greater than n. The result of this process applied to the case of secondary and ternary twinning is shown in figure 9 where the grey first-order twin (highlighted in cyan) has four secondary twins (in blue), one of which having a ternary twin (in red). Obviously, this segmentation could be argued against and one could consider that what what considered to be first-order twin is actually another grain. Unfortunately, no quantitative way to make this decision can be derived in such a case.

The indirect benefit of detecting and tagging higher order twins lies in the fact that, because of their decreasing likelihood, they help the user to check the software results more rapidly.

3.5. The particular case of "twin strips"

In a not insignificant number of cases, small twins appear as a strip of connected twins at the outcome of the previous algorithm. This phenomenon occurs either when a very thin twin is separated in small objects because of the low resolution of a few EBSD measurement points or when a twin is divided into two parts by another twin. However, it is statistically relevant to count these connected twins, also called "twin strips", as single twins.

Two connected components are considered to belong to the same twin or twin strip if they meet the following five conditions:

 i they are in the same grain;

 ii they have the same orientation, or the disorientation between the average orientation of both components is small. Typically, the same threshold is used as the one used to build connected components;
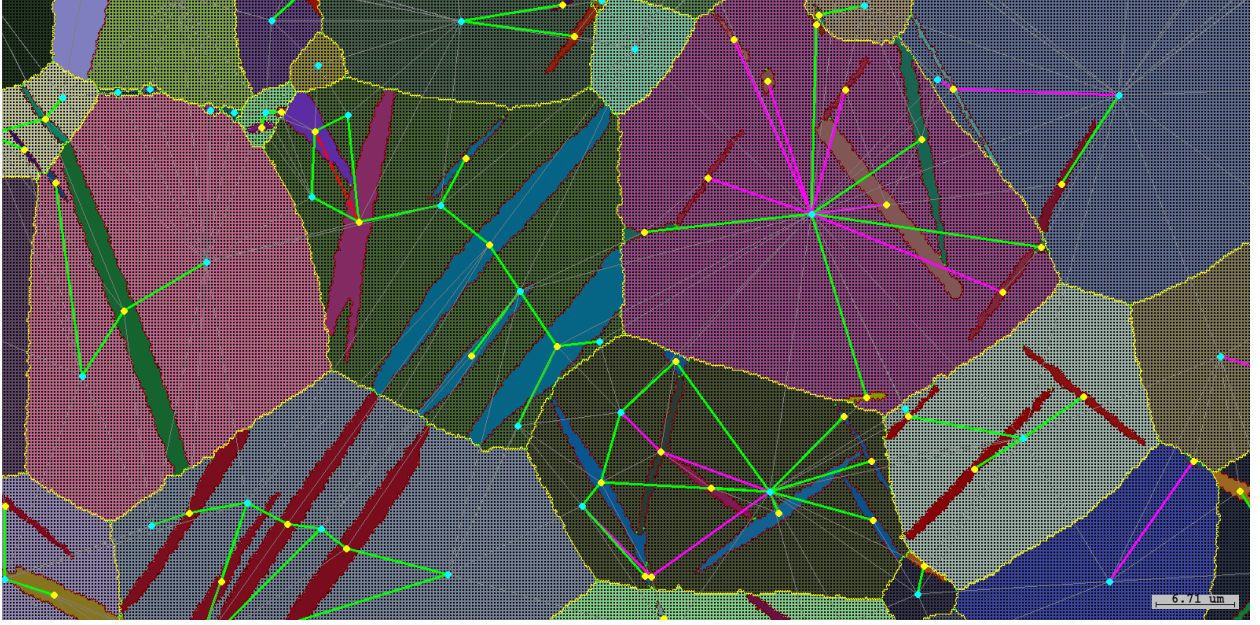
Figure 8: Zoom on the map of figure 6 to illustrate complex grain structures recovered by our software. The dashed line is a disorientation relation that match a known relation (compressive 1) but is identified as irrelevant to the twinning process.

iii the twin's ellipse main orientations (see sec. 4.4) are similar, for instance, less than 5 degrees apart;

iv the sum of the twin half-lengths are within 20% of the distance between their centroid;

v the vector linking their centroid diverges by less than a few degrees from the twin's ellipse main orientation.

From these conditions, a fifth graph is generated in all grains. Vertices are connected components and edges link pairs of connected components fulfilling the previous five conditions. Consequently, the connected components with more than one vertex are twin strips. Figure 10 gives an example of the type of reconstruction obtained with this approach.

4. Examples of automatically extracted metrics and statistics

The identification of twins as well as the computation of statistics pertaining to twinning [29] rely on metrics. The choice of these metrics and statistics was motivated by the need of a rigorous identification of twins (see sections about convexity and twin shape) but also by their known applications [19, 20, 35]. The rest of the present paragraph is dedicated to the listing and the description of some of the original software's features. These features were aimed at ensuring the correct identification of twins and easing the extraction of statistics and microstructural data via SQL requests and other post-processing codes [29]. Note that since their implementation, many other capabilities have been added to the software such as the systematic characterization of the different twin-twin junctions, secondary twins, triple junctions, etc. Regarding the twin-twin junctions, the software is now able to identify all types of twin-twin junction variants, count them, output the average disorientation angle as well as the standard deviation for any twin-twin junction variant.
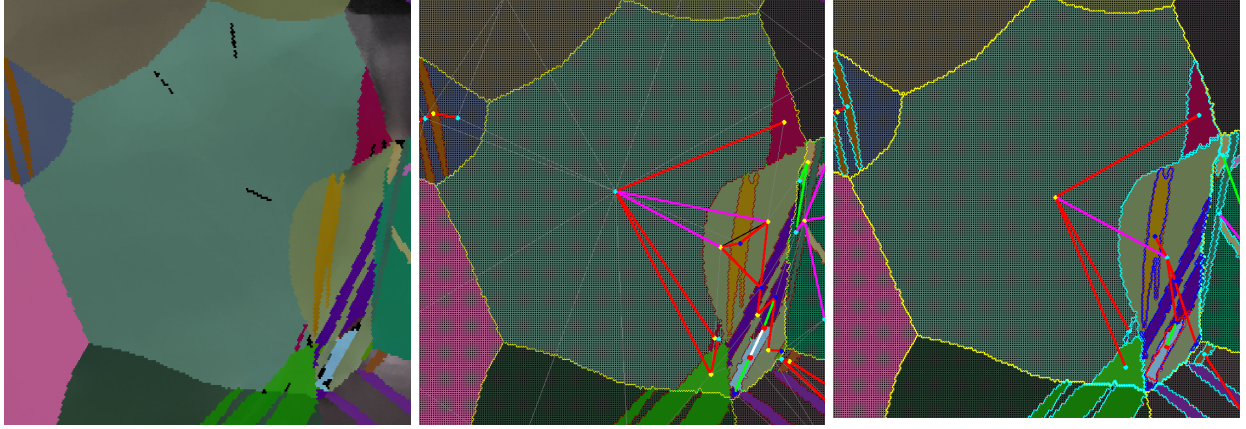
Figure 9: Example of secondary and ternary twinning observed in an EBSD map of a high-purity clock-rolled Zr sample loaded in compression along the through-thickness direction up to 3% strain. This is shown using three different visualization modes (see appendix A): raw mode (left), twinning editor mode (middle) and twinning statistics mode (right). The parent grain is surrounded in yellow, first order twins appear in cyan, secondary twins in blue and ternary or higher order twins in red.

4.1. Area and perimeter

The area of connected fragments, i.e. grains, twins, parent phases, is obtained by multiplying the number of measurement points with the area corresponding to a single pixel. The area associated with a measurement point depends on the step size and the grid type (hexagonal or square).

Grain and twin boundary lengths are computed differently. Depending on the type of measurement grid, the space between grains can be seen as a sequence of neighboring triangles or squares. The software builds such a sequence by recursively exploring the space between grains. Consider now the case of an hexagonal measurement grid - a similar discussion applies to a square grid-. Along a grain or twin boundary, the inter-fragment space is made of triangles that contain at least two measurement points belonging to the same fragment. When the three measurement points belong to different fragments, there exists a connection between either two grain boundaries or between a grain boundary and a twin boundary. These identified boundary segments are then grouped into a complete grain or twin perimeter. The software computes the grain or twin boundary length as the length of the line that links the center of all the inter-fragment triangles. The position of the contact points along the twin and grain boundaries as well as the position of the triple junctions are stored in the database.

4.2. Grain boundary properties

In order to identify grains, parent and twin phases, the disorientations between every pair or measurement points and connected parts is computed. All disorientation values are stored in a SQL database. Consequently, the user has the possibility to extract a very large amount of statistics about grain boundaries such as, for example, the distribution of disorientation angles between pixels located along the boundary, the distribution of the average disorientation angle between twin and parent phases, the average grain boundary length between two neighboring grains, etc. Dealing with one-dimensional statistics, the user is free to differentiate data related to twin boundaries, twinned and/or non-twinned grain boundaries.
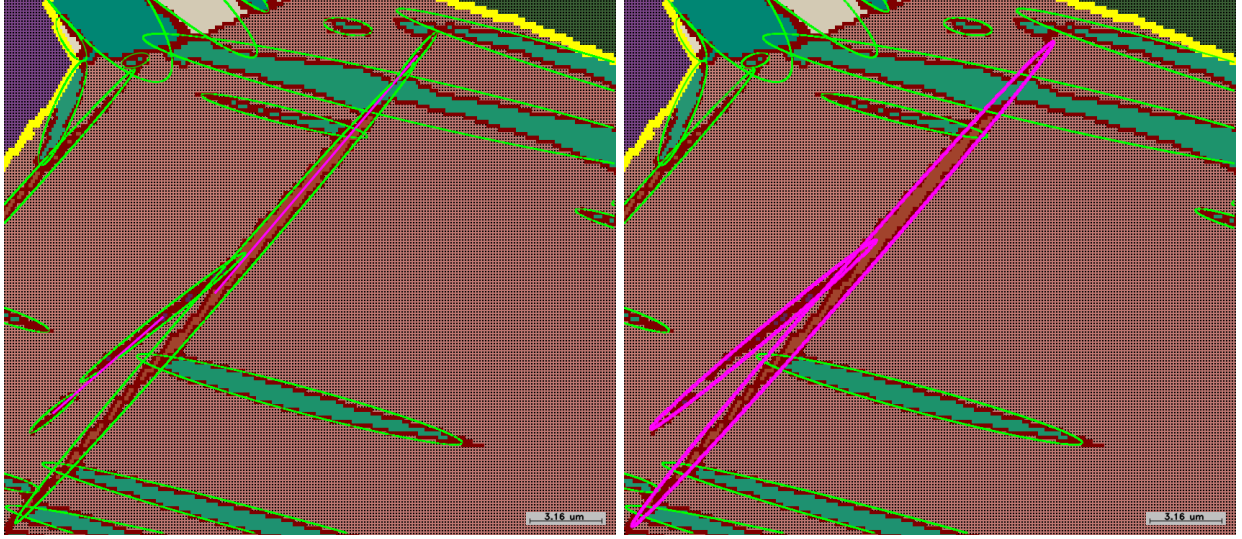
Figure 10: Zoomed-in EBSD map of a high-purity clock-rolled Zr sample loaded in compression along one of the in-plane directions up to 10% strain. Left: detected component with their ellipses and twin-strip links in magenta; right: reconstructed complete twin.

4.3. Convexity

Visual observation of EBSD maps suggests that every grain or twin seems to be more or less convex. The degree of convexity can be quantified for each item as follows. First, the convex hull of the object of interest is built by building the convex hull of all its joint points. The object can be either a grain or a twin. The convex hull is a convex polygon that encloses a set of 2D or 3D points. The construction of the convex hull can be performed in the same time as sorting the point by abscissa ($O(n \log n)$). Computing the area of such a polygon is a standard algorithm from computational geometry. The degree of convexity can then be defined as the ratio of the area of the object to the area of its corresponding convex hull. The ratio is expected to be lower than 1 and the further away from 1 it is, the less convex the object is.

This measure of convexity was implemented to refine the grain detection by identifying grains with a low convexity, trying to break edges in the connected part graph and recomputing the convexity of the resulting grains. If the overall convexity is improved, this edge is removed. Tests showed that this method is really efficient. However, the current automatic grain extraction method combined with the graphic user interface is performant enough that we rarely require the use of this refinement step.

4.4. Twin shape and ellipsicity

Length and thickness of twins are computed by estimating the 2D covariance of their constituent EBSD measurement points. The eigenvectors of the covariance matrix indicate the main directions of the twin. The apparent twin length is estimated as being equal to four times the square-root of the largest eigenvalue of its covariance matrix and the apparent twin thickness is assumed to be equal to four times the square-root of the smallest eigenvalue. The orientation of the twin main axis is then given by the orientation of the largest eigenvector of its covariance matrix.Both the true twin length and thickness can be computed in
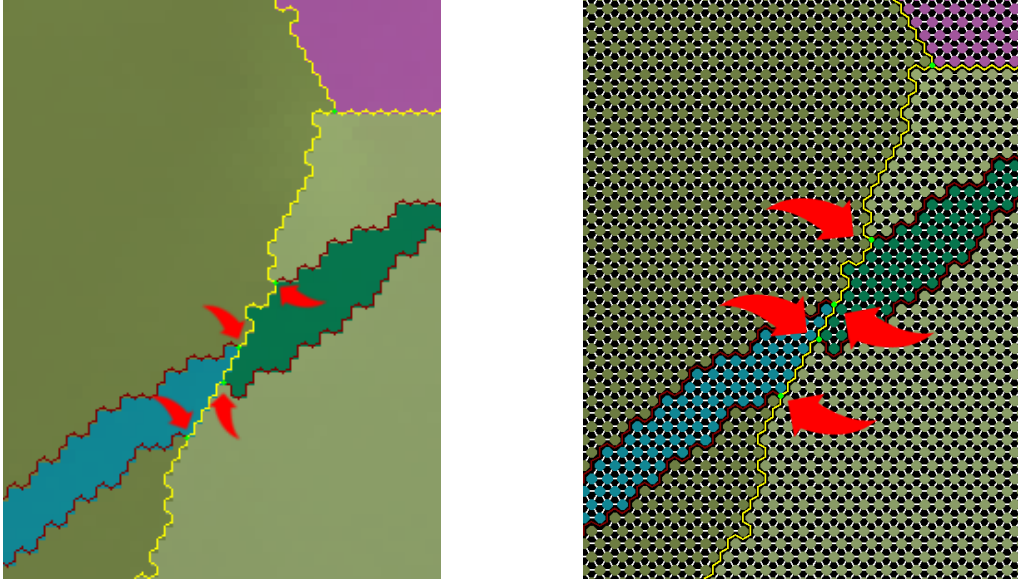
Figure 11: Zoom on an EBSD map highlighting the detection and recording of triple junctions, marked as small green dots (check the red arrows), between twins (surrounded in brown) in different grains (surrounded in yellow). The sample is high-purity clock-rolled Zr sample loaded in compression along one of the in-plane directions up 10% strain.

post-processing by multiplying them by the cosine of the angle between the twin plane and the normal to the sample [15, 20].

Comparing the ellipse area with the actual twin area computed from the number of measurement points contained in the twin allows to estimate the ellipsicity of a twin, defined as its departure from an idealized ellipse. Figure 12 depicts these ellipses and illustrates how the ellipsicity criterion highlights what the software identifies as "merged" twins. These "merged" twins are very likely co-zonal Tensile 1 twin variants. The non-distinction between co-zonal tensile 1 variants can be explained by the combination of three factors. First, the theoretical disorientation between co-zonal tensile 1 variants is equal to 9.6 degrees. Second, the software uses a 5 degree tolerance for twin variant recognition. Third, the level of precision inherent to EBSD measurements has to be considered and justifies the use of the 5 degree tolerance by the software. Consequently, reducing the software tolerance for twin recognition would not necessarily lead to a more accurate identification of two co-zonal intersecting twins but more likely to a higher number of unrecognized twins.

5. Conclusion

The present article introduces a new software for EBSD map automated analysis based on graph theory and quaternion algebra (METIS). Quaternions allow an easy computation of disorientations between pixels and areas of similar orientation. The subsequent use of graph and group structures allow grain identification, twin recognition and statistics extraction.

The newly introduced software distinguishes from pre-existing commercial softwares or academic codes by combining visualization with automated analysis of the EBSD map. The built-in graphic user interface
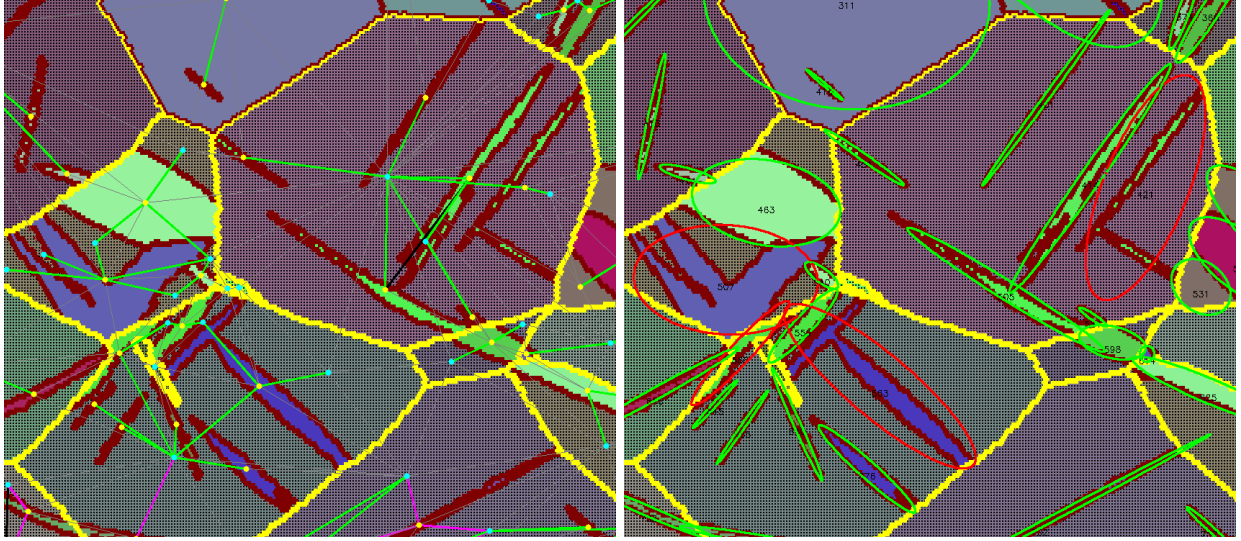
Figure 12: EBSD map of a high-purity clock-rolled Zr sample loaded in compression along one of the in-plane directions up 10% strain. The top part shows the twinning relation identified. The right caption displays ellipses fitted to twins. Red ellipses correspond to low ellipsicity (below 70%). Low ellipsicity twins correspond here to merged orthogonal twins.

enables an immediate and direct access to microstructural and twinning data such as orientation and size of twins and grains, mode and system of twins but also allows the user to correct or complete, if necessary, the analysis performed by the software. In addition, all raw and processed data are saved in a relational database. Consequently, all experimental parameters, microstructural data and twinning statistics are easily accessible via SQL requests. The database enables the systematic quantification of the influence of a very large number of parameters. The construction of such a database makes a significant difference compared to other pre-existing analysis tools.

Moreover, although initially developed to perform statistical analyses on Mg and Zr [29] scans, the software is not limited to these two h.c.p. metals. Its algorithm is capable of identifying any twin occurring in h.c.p. materials on condition that the user writes in the code the value of the c/a ratio and the theoretical disorientation quaternions corresponding to all potentially active twinning systems. For the analysis of other crystallographic structures, the user has to adapt the cell characteristics and modify the symmetry quaternions.

6. Acknowledgements

7. References

[1] P. Partridge, The crystallography and deformation modes of hexagonal close-packed metals, Metallurgical Review 12 (118) (1967) 168 – 192.

[2] M. Yoo, Slip, twinning and fracture in hexagonal close-packed metals, Metallurgical and Materials Transactions A-Physical Metallurgy and Materials Science 12 (3) (1981) 409–418.

[3] M. Yoo, J. Morris, K. Ho, S. Agnew, Nonbasal deformation modes of HCP metals and alloys: Role of dislocation source and mobility, Metallurgical and Materials Transactions A-Physical Metallurgy and Materials Science 33 (3, SI) (2002) 813–822.

[4] G. Kaschner, C. Tomé, I. Beyerlein, S. Vogel, D. Brown, R. McCabe, Role of twinning in the hardening response of zirconium during temperature reloads, Acta Materialia 54 (11) (2006) 2887 – 2896.

[5] R. McCabe, E. Cerreta, A. Misra, G. Kaschner, C. Tomé, Effects of texture, temperature and strain on the deformation modes of zirconium, Philosophical Magazine 86 (23) (2006) 3595–3611.

[6] R. J. McCabe, G. Proust, E. K. Cerreta, A. Misra, Quantitative analysis of deformation twinning in zirconium, International Journal of Plasticity 25 (3) (2009) 454 – 472.

[7] R. Cahn, Plastic deformation of alpha-uranium; twinning and slip, Acta Metallurgica 1 (1) (1953) 49 – 70.

[8] A. Crocker, The crystallography of deformation twinning in alpha-uranium, Journal of Nuclear Materials 16 (3) (1965) 306 – 326.

[9] M. Bevis, A. Crocker, Twinning shears in lattices, Proceedings of the Royal Society of London Series A-Mathematical and Physical sciences 304 (1476) (1968) 123–&.

[10] J. Christian, S. Mahajan, Deformation twinning, Progress in Materials Science 39 (1-2) (1995) 1 – 157.

[11] C. Tomé, R. Lebensohn, U. Kocks, A model for texture development dominated by deformation twinning: Application to zirconium alloys, Acta Metallurgica et Materialia 39 (11) (1991) 2667 – 2680.

[12] J. Wang, J. Hirth, C. TomTomé, Twinning nucleation mechanisms in hexagonal-close-packed crystals, Acta Materialia 57 (18) (2009) 5521 – 5530.

[13] A. Serra, D. Bacon, R. Pond, The crystallography and core structure of twinning dislocations in HCP metals, Acta Metallurgica 36 (12) (1988) 3183–3203.

[14] C. C. Aydiner, J. V. Bernier, B. Clausen, U. Lienert, C. N. Tomé, D. W. Brown, Evolution of stress in individual grains and twins in a magnesium alloy aggregate, Physical Review B 80 (2).

[15] Marshall, PE and Proust, G and Rogers, JT and McCabe, Automatic twin statistics from electron backscattered diffraction data, Journal of Microscopy 238 (3) (2010) 218–229.

[16] Mason, TA and Bingert, JF and Kaschner, GC and Wright, SI and Larsen, RJ, Advances in deformation twin characterization using electron backscattered diffraction data, Metall. Mater. Trans. A 33 (2002) 949–954.

[17] Henrie, BL and Mason, TA and Hansen, BL, A semiautomated electron backscatter diffraction technique for extracting reliable twin statistics, Metall. Mater. Trans. A 35 (2004) 3745–3751.

[18] Henrie, BL and Mason, TA and Bingert, JF, Automated twin identification technique for use with electron backscatter diffraction, Mater. Sci. Forum 495 - 497 (2005) 191–196.

[19] L. Capolungo, P. Marshall, R. McCabe, I. Beyerlein, C. Tomé, Nucleation and growth of twins in zr: A statistical study, Acta Materialia 57 (20) (2009) 6047 – 6056.

[20] I. J. Beyerlein, L. Capolungo, P. E. Marshall, R. J. McCabe, C. N. Tome, Statistical analyses of deformation twinning in magnesium, Philosophical Magazine 90 (16) (2010) 2161–2190.

[21] R. Lebensohn, C. Tomé, A self-consistent anisotropic approach for the simulation of plastic deformation and texture development of polycrystals: Application to zirconium alloys, Acta Metallurgica et Materialia 41 (9) (1993) 2611 – 2624.

[22] P. Turner, N. Christodoulou, C. Tomé, Modeling the mechanical response of rolled zircaloy-2, International Journal of Plasticity 11 (3) (1995) 251 – 265.

[23] G. Proust, C. Tomé, G. Kaschner, Modeling texture, twinning and hardening evolution during deformation of hexagonal materials, Acta Materialia 55 (6) (2007) 2137 – 2148.

[24] B. Clausen, C. Tomé, D. Brown, S. Agnew, Reorientation and stress relaxation due to twinning: Modeling and experimental characterization for mg, Acta Materialia 56 (11) (2008) 2456 – 2468.

[25] P.-A. Juan, S. Berbenni, M. Barnett, C. Tomé, L. Capolungo, A double inclusion homogenization scheme for polycrystals with hierarchal topologies: application to twinning in mg alloys, International Journal of Plasticity 60 (2014) 182 – 196.

[26] H. Wang, P. Wu, C. Tomé, J. Wang, Study of lattice strains in magnesium alloy {AZ31} based on a large strain elastic-viscoplastic self-consistent polycrystal model, International Journal of Solids and Structures 49 (15 - 16) (2012) 2155 – 2167.

[27] Dam, E.B. and Koch, M. and Lillholm, M., Quaternions, interpolation and animation, Datalogisk Institut, Kobenhavns Universitet, 1998.

[28] do Carmo, M.P., Riemannian geometry, Birkhäuser Boston, 1992.

[29] P.-A. Juan, C. Pradalier, S. Berbenni, R. J. McCabe, C. N. Tomé, L. Capolungo, A statistical analysis of the influence of microstructure and twin-twin junctions on nucleation and growth of twins in zr, Acta Macterialia 95 (2015) 399– 410.

[30] IEEE, Class segmentation and object localization with superpixel neighborhoods, Computer vision.

[31] Achanta, R. and Shaji, A. and Smith, K. and Lucchi, A. and Fua, P. and Süsstrunk, S., SLIC Superpixels compared to state-of-the-art of Superpixel Methods, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (11) (2012) 2274–2282.

[32] Bondy, A. and Murty, U.S.R., Graph Theory with Applications, North-Holland, 1976.

[33] Burtsev, S and Kuzmin, Y.P., An efficient flood-filling algorithm, Computers and Graphics 17 (5) (1993) 549–561.

[34] E. Martin, L. Capolungo, L. Jiang, J. Jonas, Variant selection during secondary twinning in mg3al, Acta Materialia 58 (11) (2010) 3970 – 3983.

[35] S. R. Niezgoda, A. K. Kanjarla, I. J. Beyerlein, C. N. Tomé, Stochastic modeling of twin nucleation in polycrystals: An application in hexagonal close-packed metals, International Journal of Plasticity 56 (0) (2014) 119 – 138.

8. Appendix A: Graphical User Interface

In addition to providing a quick and direct access to a large choice of metrics and statistics displayed on an EBSD map, the graphic user interface is also aimed at allowing the user to check and correct, if necessary, software results. Three types of situation require the intervention of the user. First, because of the random distribution of disorientations between neighboring grains, it is statistically likely that, for a few grains per map, the disorientation between adjacent grains matches a twinning relation. The user has the choice between either using the convexity option to refine the analysis performed by the map or simply deactivating manually the edge linking the parent to the mistaken twin. Secondly, in highly strained sample scans, we have observed that because of orientation gradients the disorientation between two connected parts is sometimes above the threshold to be flagged as a twinning relation. The user can manually activate this twinning relation. Lastly, the parent is by default the largest connected component of the grain. However, twin phases may occupy the largest part of the grain and appear as the parent phase. Once again, this can be corrected manually by the user. The result of such manual editions can be seen by comparing figures 6 and 7.

To assist the EBSD map analysis, 9 distinct visualization modes and 8 color mappings of the EBSD data are provided. The color mappings correspond to two different color mappings of the rotation space,

three different stereographic projections; and displayed in gray levels, the image quality, confidence index and fitness value reported by the acquisition software.

The 9 visualization modes are the following:

1. Raw mode: displays the raw EBSD measurement points as shown in figure 9, left.

2. Twinning editor: displays twinning relations between fragments and allows the user to enable or disable them. Fragments identified as parent phases, first, second, third and higher generation twins are marked with yellow, light blue, dark blue and red discs, respectively. Discs indicating high order twins, i.e. second, third and higher order, appear larger in order to be more visible. Likely being the result of incorrectly enabled relations, these twinning relations have to be inspected in priority (see figure 9, center).

3. Grain neighbors: displays grains with their neighbors. The user, here, is able to mark a fragment as parent or twin phase.

4. Clusters: displays phases grouped by twinning order. Colors used to indicate twinning modes are also the same as those used in the twinning editor visualization mode. However, it emphasizes twinning order. For example, it was previously mentioned that first generation twins are marked with a light blue disc in the twinning editor mode. Here, first generation twins are not only marked by a light blue disc but their boundaries appear in light blue (see figure 9, right).

5. Twinning statistics: exactly like the twinning editor mode except that only intra-grain relations are displayed. The interest of this mode lies in the possibility for the user to obtain statistics about twinning relations such as twinning mode, twin system activated, disorientation quaternion, etc by clicking on the edge linking the two connected parts.

6. Convex hulls: displays the convex hull of detected grains. The polygons are drawn in green if their area is close to the enclosed grain area, in red otherwise.

7. Ellipses: displays a fitted ellipse around every detected twins. Twins whose shape does not fit an ellipse very well are drawn in red because they are likely to be the result of the merging of two (or more) twins (see figure 10, left). Note that when this situation happens, merging twins belong to the same mode and system of twinning. This mode allows the user to have access to grain and twin properties such as orientation and size.

8. Connected twins: displays detected twin-twin junctions. The user can also mark manually undetected twin-twin junctions..

9. Twin joints: displays identified twinning relations between measurement points located on twin joints. Even though measurement points on joints are not very reliable, this mode might be useful to visualize how strong the disorientation appears along the boundary.

In addition to these visualization modes, options are available to highlight grain or twin joints, exclude grains in contact with the map edge, replace twins in a twin strip by their union (see figure 10, right), display

connected part ids and zoom in or out and pan. When zooming in to a level where individual measurement can be separated, local disorientation are also displayed as shown in figure 3. Figure 13 summarizes all available modes.

9. Appendix B: Data availability

The analysis of an EBSD map generates a wealth of quantitative data about grains, twins and parent phases. However, different studies will not use the same EBSD data for the same purposes. This is why it is really important to export data in a way that preserves all relations and does not make assumptions on what should or should not be stored.

The present software exports the data-structure extracted from the EBSD map analysis inside an SQL database with the structure described in figure 14. In practice, this is performed using the SQLite library that implements a server-less database stored inside a single file. The advantage of such a database is that it keeps all the information in a single file and allows the user to create aggregated statistics with simple SQL requests, as shown hereafter with requests 1 and 2. For example, Request 1 generates a table containing features about twins such as twinning modes (i.e. "twinning"), twin systems (i.e. "variant"), apparent twin area (i.e. "area"), apparent twin thickness (i.e. "thickness"), quaternions corresponding to the average twin orientation (i.e. "qx", "qy", "qz", "qw"), etc. Request 2 was used to extract informations about twin-twin junctions such as the modes and systems of intersecting twins. This request relies on the view created by Request 1 to generate the table containing twin characteristics. Matlab, C++ or Fortran codes can then process the tables generated by the SQL requests in order to extract statistics of interest. For example, statistics about the influence of microstructure and twin-twin junctions on nucleation and growth of twins presented by Juan et al. [29] results from the analysis of only four tables. Moreover, for the sake of keeping record of experimental conditions, the database also stores constants and parameters used to construct this particular EBSD analysis.

- Request 1:

```
select C2.id, C2.grain, S2.area, S2.length, S2.thickness, E.twinning,
    E.variant, S2.qx, S2.qy, S2.qz, S2.qw
from FragmentEdges as E
    inner join Fragment as C1 on C1.id = E.i
    inner join Fragment as C2 on C2.id = E.j
    inner join FragmentStats as S2 on S2.id = C2.id
    inner join Grains as G on G.id = C2.grain
where C1.is_parent and C1.grain = C2.grain
    and C1.twinstrip <= 0 and C2.twinstrip <= 0
    and C2.twinning_order = 1
    and E.twinning and not G.map_edge
group by C2.id order by C2.id;
```

- Request 2:

```
select C1.grain, C1.id, C2.id,
    T1.twinning, T1.variant, T2.twinning, T2.variant
```
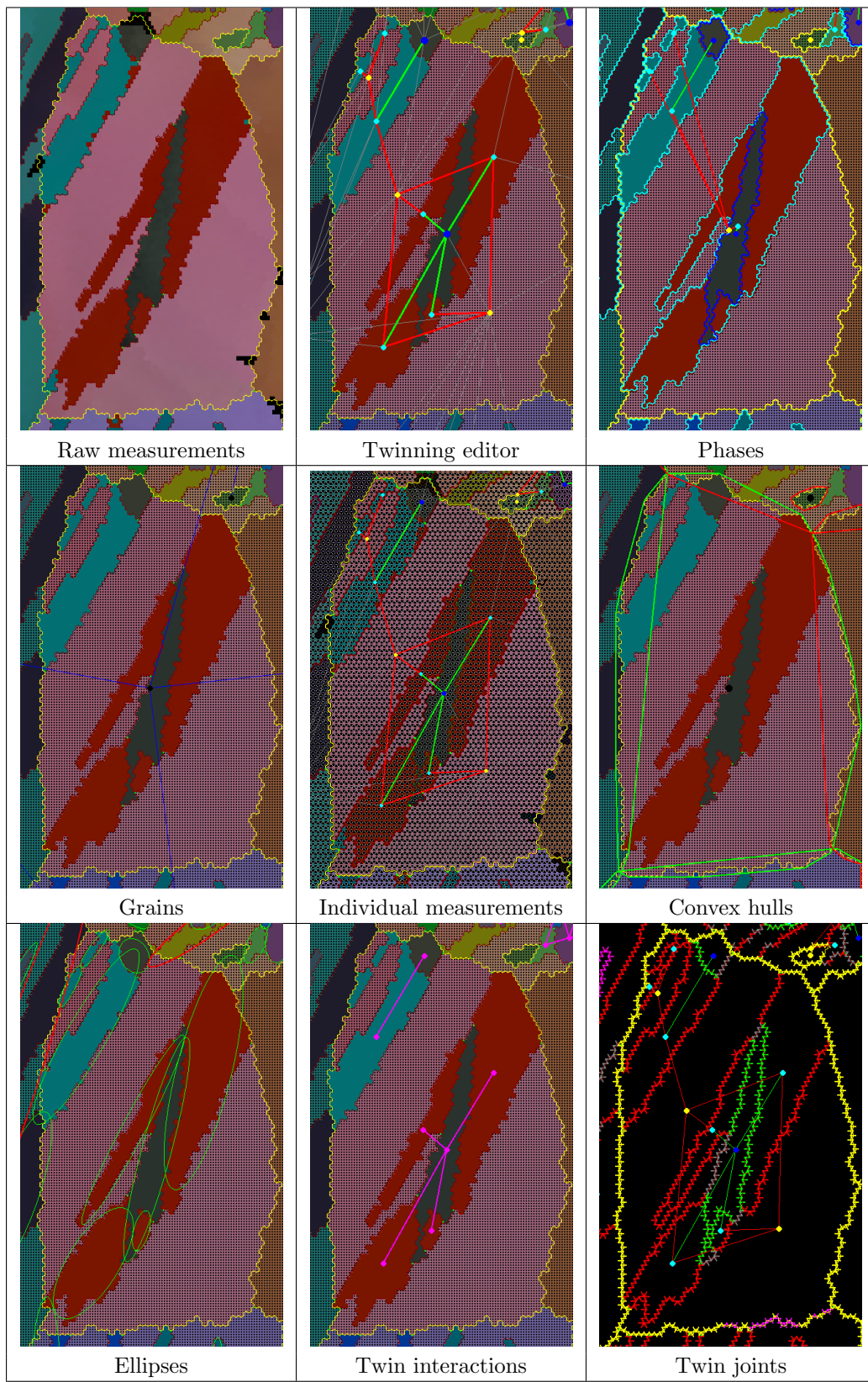
Figure 13: Visualization modes for a single grain

Grain Edges

i,j
qx,qy,qz,qw

Grains

Id

Clusters

Id
grain

2 — 0–n

1..n — 1

1..n

1..n

1

Strings

Value
Text

Phases

Id
grain

Fragments

Id
grain, phase, cluster
is_parent
twinning
twinstrip

1 — 1

1..n — 1

0..n

1
1

Constants

Name
Value

Fragment Edges

i,j
twinning
variant
qx,qy,qz,qw

Fragment Statistics
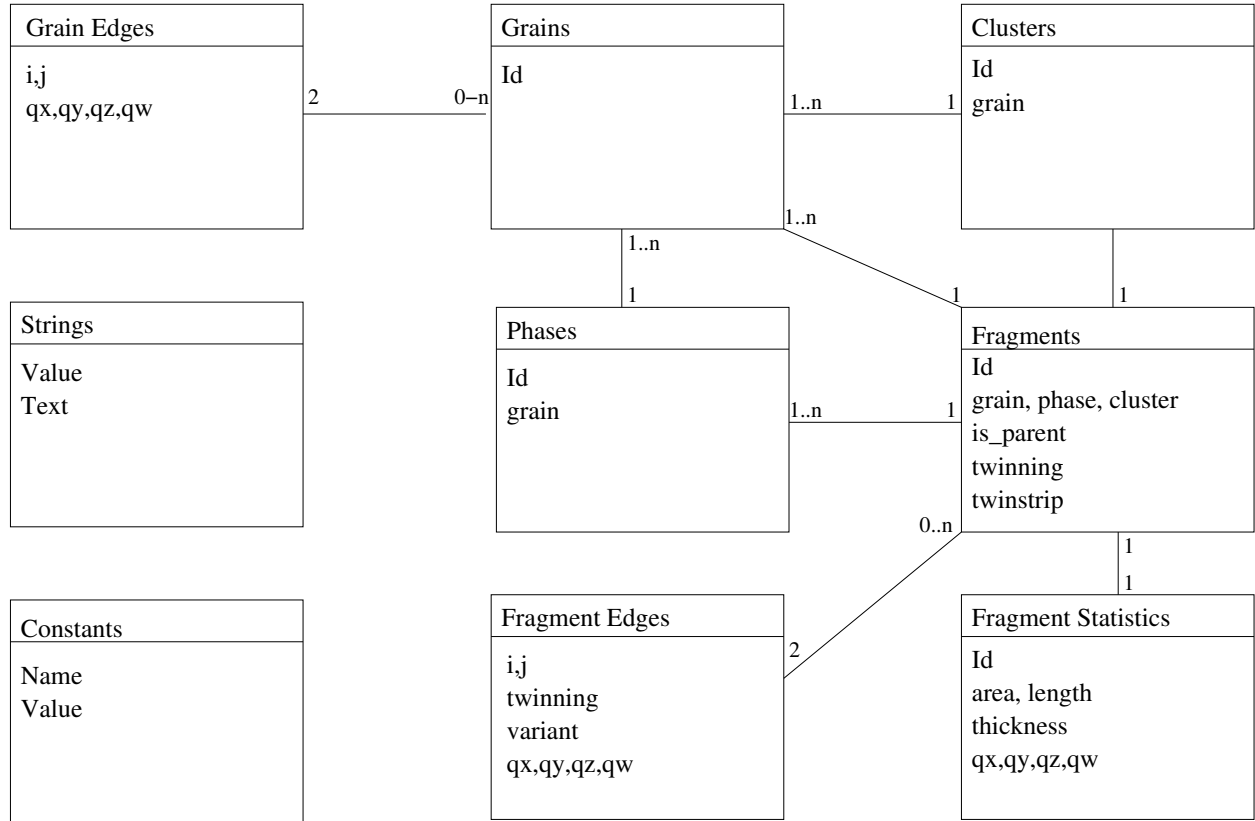
Id
area, length
thickness
qx,qy,qz,qw

2

1
1

Figure 14: Structure of the database used to store the EBSD analysis results. Boxes are database tables, edges with numbers indicates relations and the n-arity of these relations.

```
from FragmentEdges as E
    inner join Fragment as C1 on C1.id = E.i
    inner join Fragment as C2 on C2.id = E.j
    inner join Twins as T1 on C1.id = T1.id
    inner join Twins as T2 on C2.id = T2.id
where E.i > E.j
    and not C1.is_parent and not C2.is_parent
    and C1.is_valid and C2.is_valid
    and C1.grain=C2.grain
    and C1.twinstrip <= 0 and C2.twinstrip <= 0
order by C1.grain;
```